The Dr. Theo Schöller Chair of Technology and Innovation Management offers a Master’s Thesis with the following topic:

Promises and Perils of Utilizing GitHub for Management Research

Background and motivation
With Open Source Software (OSS) gaining more and more popularity among technology companies, research in the area of how companies engage in OSS development is gaining further importance. One such area of research is the field of technology acquisitions, where the acquisition of Red Hat by IBM for $34 billion at the end of 2018 just marked the tip of the iceberg of a recent acquisition spree.

Yet, quantitative data on OSS activities of companies is scarce. Here, GitHub - the largest OSS development platform globally - offers a promising resource for management research as most companies active in the OSS space use GitHub as their means of choice to open source their OSS projects or contribute to other projects, as well as GitHub’s extensive set of APIs through which the activities of these companies can potentially be tracked.

However, management research conducted based on GitHub is scarce (e.g. Nagle, 2019) and several obstacles & risks associated with analysis on firm level need to be understood (e.g. identification on employees and which of their activities are conducted as part of their job or in their spare time; see e.g. Kalliamvakou et al., 2014).

Objectives and Methodology
To gain a better understanding on how GitHub can best be utilized for firm-level analysis and which risks exist this thesis should achieve the following objectives:

▪ Analyze existing management and computer science publications based on GitHub data, with regards to focus of analysis, level of analysis (e.g. individual, firm, ecosystem), data source (e.g. GHTorrent, GH Rest API), identification of employees and corporate contributions, etc.
▪ Identify the key obstacles and risks which need to be understood to utilize GitHub for firm level analysis
▪ Describe a process how to extract the relevant data (previous work exists which might need some minor adjustments)
▪ Conduct quantitative analysis to test the usability of GitHub data regarding if and how the obstacles and risks defined before can be overcome
▪ Ideally result of the thesis would be a paper which can be presented at e.g. www.msrconf.org

What we offer
▪ Work on a topic which is highly relevant for science and practice
▪ Close mentoring and clear milestones
▪ Extensive previous work which can be utilized (literature, code and data)
▪ A fun but also challenging working environment

What we expect
▪ Interest in playing around with data/statistics is crucial for this thesis
▪ Open and pro-active communication style
▪ Independent & diligent working style
▪ Very good grades

Contact
Please contact Michael Vetter (Michael.vetter@tum.de) if you are interested in this topic. Your application should include a short letter of motivation, your CV, and a current transcript of records. For more information on our general requirements, the application procedure, the application deadlines, and the style guidelines please go to https://www.tim.wi.tum.de/index.php?id=210.